CATALOGIC MODELS FOR PREDICTING BIODEGRADATION OF CHEMICALS

D. Nedelcheva[a], D. Georgieva[a], N. Dimitrova[a], S. Dimitrov[a], O. Mekenyan[a]

[a]*Laboratory of Mathematical Chemistry, University "Prof. As. Zlatarov", 1 Prof. Yakimov St, 8010 Bourgas, Bulgaria*

INTRODUCTION

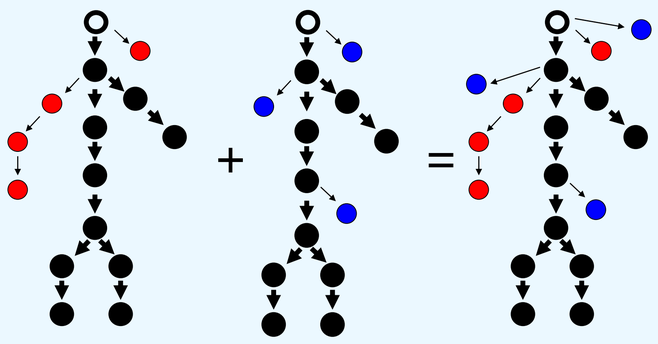
The unprecedented pollution of environment by xenobiotic compounds has provoked the need of understanding and predicting biodegradability of chemicals. Mechanistic understanding of microbial degradation is a premise for adequate modeling of environmental fate of chemicals. The aim of the presentation is to describe biodegradation models implemented in CATALOGIC software. These are principally new models based on unique mathematical formalism [Dimitrov, S.; Pavlov, T, Veith, G, Mekenyan, O SAR and QSAR in Environ Res, 22, 2011, 699-718]. Based on simulated pathways of degradation the models are able to predict biological oxygen demand (BOD), CO_2-production, primary and ultimate half-lives and quantities of transformation products [Dimitrov, S.; Pavlov, T.; Dimitrova, N.; Georgieva, D.; Nedelcheva, D.; Kesova, A.; Vasilev, R.; Mekenyan, O SAR and QSAR in Environ. Res., 22, 2011, pp. 719-755].

MATERIALS AND METHODS

The CATALOGIC software suite includes electronic databases of observed biodegradability and metabolic information. The most specific feature of CATALOGIC models is that simulated metabolism is used to extract the material balance equations witch are used to predict the consumption of oxygen, CO2 production, half-lives and quantities of metabolites. The quality of the simulators is quantified by the degree of reproducibility between observed and generated metabolites and the degree of correctly classified as ready/not ready degradable chemicals.

Observed and predicted metabolism

Union of metabolisms



● - Observed and predicted metabolites,

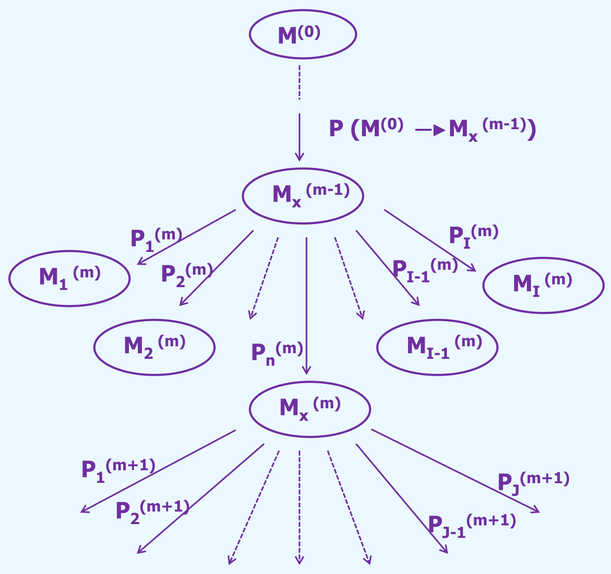
● - Observed and not predicted metabolites,

● - Predicted and not observed metabolites,

$$\text{Predictability} = \frac{\bullet}{\bullet + \bullet} \qquad \text{Sensitivity} = \frac{\bullet}{\bullet + \bullet}$$

Mathematical formalism

Generation of metabolic maps



M_x - generated metabolite P_n - Intrinsic probability

Model parameterizations

Static model:

$$\min_{\mathbf{P}} RSS = \sum_{n=1}^{N} \left(E_n^{Obs} - E_n^{Calc}(\mathbf{P})\right)^2$$

E_n= *BOD, ThCO*$_2$

P_i - intrinsic probability

$$P_i = 1 - \exp(-k_i t)$$

k_i - first order rate constant

Kinetic model:

$$\min_{\mathbf{k}} RSS = \sum_{n=1}^{N} \left(\sum_{t=1}^{T_E} \left(E_{n,t}^{Obs} - E_{n,t}^{Calc}(\mathbf{k})\right)^2 + \left(t_{1/2,n}^{Obs} - t_{1/2,n}^{Calc}(\mathbf{k})\right)^2 \right)$$

Applicability domain

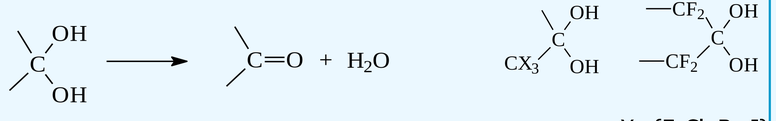
Stepwise approach

- General requirements – MW, S_W, log K_{OW}, etc.
- Structural domain – based on atom-centered fragments
- Domain of simulator of metabolism – based on reliability of simulated metabolism

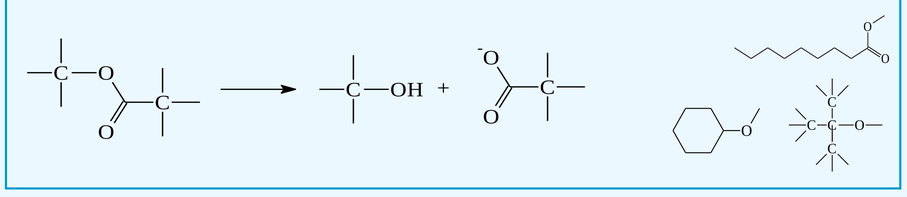
Simulation of microbial metabolism

Based on observed metabolic pathways and expert knowledge the principle catabolic transformations were determined and used to mimic microbial metabolism. Each transformation consists of source and product fragments. The application of the transformations is controlled by general (valid for all transformations) and local (valid only for one transformation) rules. These rules reflect the logic of metabolism and prevent stochastic propagation of the generated metabolic trees. Hierarchy and transformation probabilities or kinetic constants were iteratively parameterized on the basis of observed metabolites and biodegradability of training chemicals.

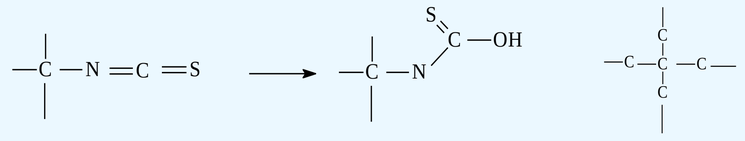
Geminal diol decomposition P= 0.99 — Inhibiting fragments



Ester hydrolysis P= 0.80 — Inhibiting fragments



Isothiocyanate hydrolysis P= 0.26 — Inhibiting fragments



RESULTS AND DISCUSSION

Biodegradation OECD 301 C model

The correctness of classification between observed and predicted ready/not ready chemicals (*BOD* < 50% or *BOD* > 70%) is 96% and 91%, respectively.

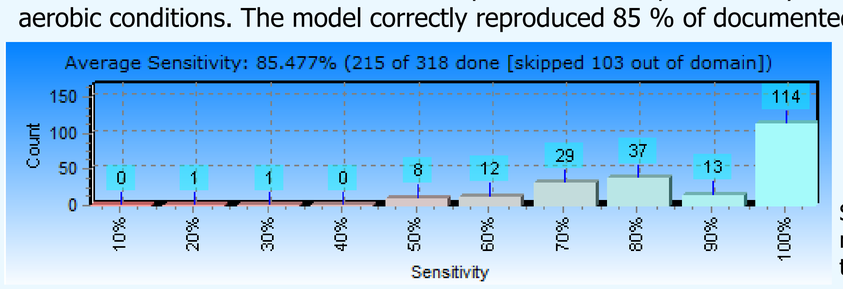
OECD 301C model C* = 0.77, C* = 0.91[1]		Observed biodegradation: Not ready BOD < 50	Border range 50 =< BOD =< 70	Ready BOD > 70
Predicted biodegradation	Not ready BOD < 50	638	15	30
	Border range 50 =< BOD =< 70	33	29	46
	Ready BOD > 70	24	58	205
	Total	695	102	281

(*C**) -adjusted Pearson's contingency coefficient

[1] - Observed or predicted *BOD* from the range 50 =< *BOD* =< 70 are not accounted for.

Soil degradation model

The Soil BioPath model is focused on reproduction of experimentally observed metabolites in terrestrial aerobic conditions. The model correctly reproduced 85 % of documented metabolic maps.

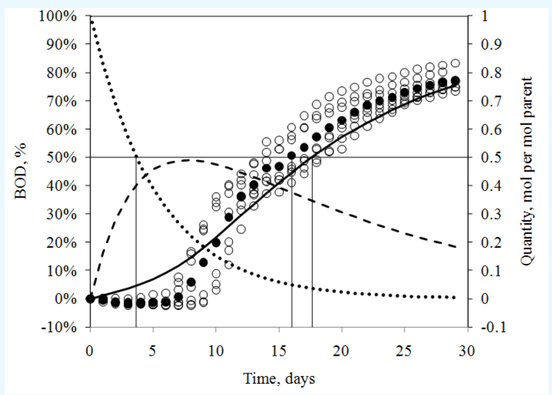


Sensitivity - probability that the metabolite is predicted, given that the metabolite is truly observed

Biodegradation OECD 301 F model

Model predictions:

- variation of quantities of 2-Methylpyridine and its first metabolite
- primary and ultimate half-lives



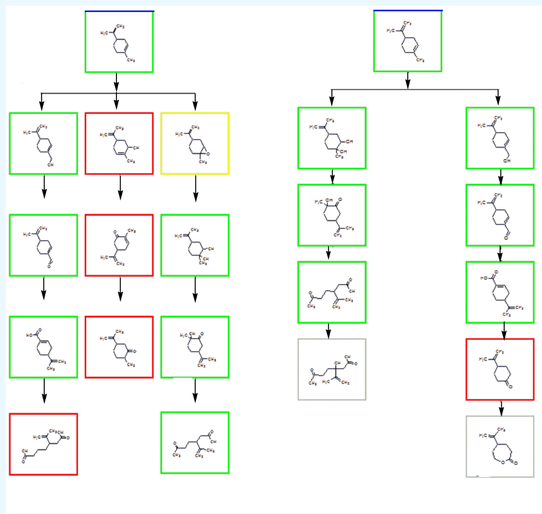
○ - observed BOD
● - mean of observed BOD
—— - predicted BOD
····· - predicted quantity of 2-Methylpyridine
- - - - predicted quantity of the first metabolite

predicted primary half-life $t_{1/2}^{\text{Pr Calc}} = 4\,\text{days}$

observed ultimate half-live $t_{1/2}^{\text{Ult Obs}} = 16\,\text{days}$

predicted ultimate half-life $t_{1/2}^{\text{Ult Calc}} = 18\,\text{days}$

Comparison of observed and simulated metabolism of Limonene



The predictions made in CATALOGIC can be generalize in the summary report

CAS Name SMILES	BOD at 28th day,	95% confidence interval	Q[Parent] at 28th day*	Half-lives, days: Primary	Half-lives, days: Ultimate	10 days window BOD	Sub-Domains: General requirements	Sub-Domains: Structural	Sub-Domains: Metabolic	Summary domain
109-06-8 2-methylpyridin c1(C)ccccn1	74%	± 210%	0.0051	3.6	17	48% Not ready	In sub-domain	In sub-domain	In sub-domain	In domain

SUMMARY AND CONCLUSIONS

1. The CATALOGIC 301C and 301F models correctly classify more than 80% and 90% of training chemicals, respectively.
2. The Soil BioPath model which is focused on reproduction of experimentally observed metabolites in soil correctly predicts 85% of documented metabolites.
3. Biodegradation models presented here meet the requirements (mechanistic interpretation, defined endpoint, scientific validity, applicability domain, documentation, etc.) necessary to replace testing for determining degradability of chemicals identifying the absence or presence of stable metabolites.

The poster can be downloaded at: **http://oasis-lmc.org/posters/QSAR2012**